



ELECTRONIC FRONTIER FOUNDATION

Home » Deeplinks Blog » October, 2008

October 31st, 2008

Google Book Search Settlement: A Reader's Guide

Legal Analysis by Fred von Lohmann



UPDATE (Aug. 2009): If you are an **author**, head over to our [Google Books page](#).

UPDATE (Jan. 2009): The official [class notice](#) has now been published. Anyone who owns a copyright and has questions about the settlement should start there. Also, I strongly recommend Prof. James Grimmelmann's

[analysis](#) of the settlement.

As we [reported](#) earlier this week, Google has settled the lawsuit brought in 2005 by authors and book publishers regarding its massive book scanning and indexing project. Although the settlement must still be approved by the court and is unlikely to go into effect until sometime late in 2009, commentary has already been flooding the blogosphere. Generally, opinions are split between excitement for users ("better access to zillions of out-of-print books") and suspicion of Google ("one library to rule them all, and in the darkness bind them").

We are still digesting the ~300-page proposed [settlement agreement](#) (for those seeking a good overview, the [39-page notice](#) to class members is a good place to start).

So far, two things are plain.

First, this agreement is likely to change forever the way that we find and browse for books, particularly out-of-print books. Google has already scanned more than 7 million books, and plans to scan millions more. This agreement will allow Google to get close to its original goal of including all of those books into Google's search results (publishers got some concessions, however, for in-print books). In addition to search, scanned public domain books will be available for free PDF download (as they are today). But the agreement goes beyond Google's Book Search by permitting access, as well. Unless authors specifically opt out, books that are out-of-print but still copyrighted will be available for "preview" (a few pages) for free, and for full access for a fee. In-print books will be available for access only if rightsholders affirmatively opt in. The upshot: Google users will have an unprecedented ability to search (for free) and access (for a fee) books that formerly lived only in university libraries.

Second, this outcome is plainly second-best from the point of view of those who believe Google would have won the [fair use question](#) at the heart of the case. A legal ruling that scanning books to provide indexing and search is a fair use would have benefited the public by setting a precedent on which everyone could rely, thus limiting publishers' control over the activities of future book scanners. In contrast, only Google gets to rely on this settlement agreement, and the agreement embodies many concessions that a fair user shouldn't have to make.

But the settlement has one distinct advantage over a litigation victory: it's much, much faster. A complete victory for Google in this case was probably years away. More importantly, a victory would only have given the green light for scanning in order to index and provide snippets in search results; it would not have provided clear answers for all the other activities addressed in the settlement, such as providing display access for out-of-print books, allowing nondisplay research on the corpus, and providing access for libraries. Litigating all of those fair use questions could easily have taken a decade or more. As

University of Michigan head librarian Paul Courant points out, those are years that we would never get back. (University of Virginia's Prof. Siva Vaidhyanathan offers a differing view: "These claims are not convincing when one considers just how great an alternative system could be, if everyone would just mount a long-term, global campaign for it rather than settle for the quick fix.").

Conclusions beyond those two are harder to draw. Many devils are buried in the details of the 300-pages of legalese, and much will turn on how the agreement is implemented. Here are the 6 "big picture" concerns that I'm keeping in mind as I review those details:

Fair Use: How will this agreement impact future fair use cases involving book scanning? Others (like the Open Content Alliance) are scanning books, and they may not have Google's ability (or budget) to strike a deal with the world's publishers. UCLA Law's Prof. Neal Netanel has a few preliminary thoughts along this line at the Balkinization blog.

Innovation: It seems likely that the "nondisplay uses" of Google's scanned corpus of text will end up being far more important than anything else in the agreement. Imagine the kinds of things that data mining all the world's books might let Google's engineers build: automated translation, optical character recognition, voice recognition algorithms. And those are just the things we can think of today. Under the agreement, Google has unrestricted, royalty-free access to this corpus. The agreement gives libraries their own copy of the corpus, and allows them to make it available to "certified" researchers for "nonconsumptive" research, but will that be enough?

Competition: In the words of Prof. Michael Madison, "Has Google backed away from an interesting and socially constructive fair use fight in order to secure market power for itself?" Does this deal give Google an unfair head start against any second-comers to book scanning? The agreement creates an independent, nonprofit Book Rights Registry to dole out Google's royalties, and the parties clearly hope that the Registry will be able to license others on similar terms. But the Registry is empowered to cut a deal with Google on behalf of *all* rightsholders by virtue of the class action; in order to offer similar blanket licenses to others, it would have to independently acquire rights from each and every copyright owner individually. How long will that take? What about the Registry itself? It hopes to be a monopoly that fixes prices for the entire market of copyright owners -- precisely the kind of thing that landed ASCAP and BMI, which dole out blanket licenses for music, in antitrust trouble decades ago.

Access: This agreement promises unprecedented access to copyrighted books. But by settling for this amount of access, has Google made it effectively impossible to get *more* and better access? The agreement allows you to "purchase" digital access for out-of-print books, but does not include the right to download the book (unlike public domain books). So you can read the book, but only on Google's terms. Libraries get more access, but for an undisclosed price (OK, one computer for free) and still with a variety of restrictions. In the words of Harvard's head librarian, "As we understand it, the settlement contains too many potential limitations on access to and use of the books by members of the higher education community and by patrons of public libraries."

Public Domain: Early reports are that public domain materials are not regulated by the agreement. Moreover, Google has negotiated a "safe harbor" that protects it from liability for mistakes in evaluating the copyright status of a book. That should result in more willingness to forge ahead with the free PDF posting of books published between 1923-1963, where a public domain determination turns on checking government records to see whether the copyright had been renewed. But will Google impose restrictions on these "safe harbor" public domain works? Will the libraries that receive a digital copy of their own public domain holdings impose restrictions on those copies?

Privacy: The agreement apparently envisions a world where Google keeps all of the electronic books that you "purchase" on an "electronic shelf" for you. In other words, in order to read the books you've paid for, you have to log into Google. Google is also likely to keep track of which books you browse (at least if you're logged in). This is a huge change in the privacy we traditionally enjoy in libraries and bookstores, where nobody writes down "Fred von Lohmann entered the store at 19:42:08 and spent 2.2 minutes on page 28 of 0-486-66980-7, 3.1 minutes on page 29, and 2.8 minutes on page 30." If Google

becomes the default place to search, browse, and buy books, it will be able to keep unprecedented track of what you read, how you read it, and collate that with all the other information it has about you. Does the agreement contain ironclad protections for user privacy?

Related Issues: [Innovation](#), [Intellectual Property](#)

Related Cases: [Author's Guild v. Google](#)

[Permalink: <http://www.eff.org/deeplinks/2008/10/google-books-settlement-readers-guide>] [[Email this article](#)]

Printed Material Notice: Any and all original material on the EFF website may be freely distributed at will under the Creative Commons Attribution-NonCommercial License, unless otherwise noted. All material that is not original to EFF may require permission from the copyright holder to redistribute.